



TITLE:

Semidefinite Programming Relaxation for Nonconvex Quadratic Programs(Discrete and Continuous Structures in Optimization)

AUTHOR(S):

Fujie, Tetsuya; Kojima, Masakazu

CITATION:

Fujie, Tetsuya ...[et al]. Semidefinite Programming Relaxation for Nonconvex Quadratic Programs(Discrete and Continuous Structures in Optimization). 数理解析研究所講究録 1996, 945: 57-67

ISSUE DATE:

1996-04

URL:

<http://hdl.handle.net/2433/60225>

RIGHT:

Semidefinite Programming Relaxation for Nonconvex Quadratic Programs

東京工業大学 藤江 哲也 (Tetsuya Fujie)

東京工業大学 小島 政和 (Masakazu Kojima)

Abstract. Any quadratic inequality in the n -dimensional Euclidean space can be relaxed into a linear matrix inequality in $(1+n) \times (1+n)$ symmetric matrices. Based on this principle, we extend the Lovász-Schrijver SDP (semidefinite programming) relaxation developed for a 0-1 integer program to a general nonconvex QP (quadratic program), and present some fundamental characterization of the SDP relaxation including its equivalence to a relaxation using convex-quadratic valid inequalities for the feasible region of the QP.

Key words. Semidefinite Program, Relaxation Method, Interior-Point Method, Linear Matrix Inequality, Nonconvex Quadratic Program

1 Introduction.

We use the symbols $\mathcal{S}(m)$ for the set of $m \times m$ symmetric matrices, and $\mathcal{S}(m)_+$ (or $\mathcal{S}(m)_{++}$) for the cone consisting of $m \times m$ symmetric positive semidefinite (or positive definite, respectively) matrices. We are concerned with a canonical form QP:

$$\text{Minimize } \mathbf{c}^T \mathbf{y} \text{ subject to } \mathbf{y} \in \mathcal{F}. \quad (1)$$

Here

$$\left. \begin{aligned} \mathcal{F} &\equiv \left\{ \mathbf{y} \in R^{1+n} : y_0 = 1, \mathbf{y}^T \mathbf{P}_k \mathbf{y} \leq 0 \ (k = 1, 2, \dots, m) \right\}, \\ \mathbf{c} &\equiv \begin{pmatrix} \gamma \\ \mathbf{d} \end{pmatrix} \in R^{1+n}, \ \mathbf{y} \equiv \begin{pmatrix} y_0 \\ y_1 \\ \cdot \\ y_n \end{pmatrix} \in R^{1+n}, \\ \mathbf{P}_k &\equiv \begin{pmatrix} \pi_k & \mathbf{q}_k^T/2 \\ \mathbf{q}_k/2 & \mathbf{Q}_k \end{pmatrix} \in \mathcal{S}(1+n) \ (k = 1, 2, \dots, m), \\ \pi_k &\in R, \ \mathbf{q}_k \in R^n, \ \mathbf{Q}_k \in \mathcal{S}(n) \ (k = 1, 2, \dots, m). \end{aligned} \right\} \quad (2)$$

Note that the feasible region \mathcal{F} is contained in the n -dimensional hyperplane $H \equiv \{\mathbf{y} \in R^{1+n} : y_0 = 1\}$, and that the function $H \ni \mathbf{y} \rightarrow \mathbf{y}^T \mathbf{P}_k \mathbf{y} \in R$ involved in the inequality constraint is convex (or linear) if and only if $\mathbf{Q}_k \in \mathcal{S}(n)_+$ (or $\mathbf{Q}_k = \mathbf{O}$, respectively). \mathbf{Q}_k can be indefinite, so that the feasible region \mathcal{F} of the QP (1) is a nonconvex subset of the hyperplane H in general. The canonical form QP (1) covers various mathematical programs such as 0-1 IPs, general nonconvex QPs and linear complementarity problems.

SDP(Semidefinite Program) has been considered as a powerful tool for relaxation of many combinatorial optimization problems, since it yields a tight bound [1, 2, 5, 6, 8, 13, 17] and it is efficiently solvable by interior methods [1, 2, 3, 4, 7, 9, 12, 16]. SDP relaxation is originally proposed by Lovász[10] for stable set problems.

Among many literatures related on SDP relaxation, this paper was motivated by Alizadeh [1] where an elementary outline of the SDP relaxation method proposed by Lovász-Schrijver [11] for 0-1 IPs (integer programs) was presented. The aim of this paper is to present a general method for constructing an SDP which serves as a relaxation of the QP (1) and some fundamental properties on the SDP relaxation. Our SDP relaxation method may be regarded as a straightforward extension of the Lovász-Schrijver SDP relaxation method [11] for 0-1 IPs to the QP (1). It is also characterized in terms of

- a dual of Shor's relaxation method [14] (see also [15]) for general nonconvex QPs (this will be discussed in Section 5), and
- a relaxation using *convex-quadratic valid inequalities* for the feasible region \mathcal{F} .

Let

$$\mathbf{P} \equiv \begin{pmatrix} \pi & \mathbf{q}^T/2 \\ \mathbf{q}/2 & \mathbf{Q} \end{pmatrix} \in \mathcal{S}(1+n), \quad \pi \in \mathbb{R}, \quad \mathbf{q} \in \mathbb{R}^n, \quad \mathbf{Q} \in \mathcal{S}(n).$$

We say that an inequality $\mathbf{y}^T \mathbf{P} \mathbf{y} \leq 0$ is a convex-quadratic (or linear) valid inequality for \mathcal{F} if

$$\mathbf{Q} \in \mathcal{S}(n)_+ \text{ (or } \mathbf{Q} = \mathbf{O}, \text{ respectively) and } \mathbf{y}^T \mathbf{P} \mathbf{y} \leq 0 \text{ for every } \mathbf{y} \in \mathcal{F}.$$

Then $\text{co } \mathcal{F}$, the convex hull of \mathcal{F} is completely determined by all the convex-quadratic valid inequalities for \mathcal{F} ;

$$\text{co } \mathcal{F} = \bigcap_{\mathbf{P} \in \mathcal{V}} \{ \mathbf{y} \in \mathbb{R}^{1+n} : y_0 = 1, \mathbf{y}^T \mathbf{P} \mathbf{y} \leq 0 \},$$

where \mathcal{V} denotes the set of all matrices $\mathbf{P} \in \mathcal{S}(1+n)$ that induce convex-quadratic valid inequalities for \mathcal{F} . (The identity above is well-known when \mathcal{V} is the set of all matrices $\mathbf{P} \in \mathcal{S}(1+n)$ that induce linear valid inequalities for \mathcal{F}). The discussion above leads us to a relaxation of the QP (1) using all convex-quadratic valid inequalities for \mathcal{F} that we can generate as a nonnegative combination of the quadratic inequalities of the QP (1):

$$\text{Minimize } \mathbf{c}^T \mathbf{y} \text{ subject to } \mathbf{y} \in \tilde{\mathcal{F}}, \quad (3)$$

where

$$\left. \begin{aligned} \tilde{T} &\equiv \{ \mathbf{t} \in R^m : \mathbf{t} \geq \mathbf{0}, \sum_{k=1}^m t_k \mathbf{Q}_k \in \mathcal{S}(n)_+ \}, \\ f_{\mathbf{t}}(\mathbf{y}) &\equiv \mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \mathbf{y} \text{ for every } \mathbf{y} \in R^{1+n} \text{ } (\mathbf{t} \in \tilde{T}), \\ \tilde{\mathcal{F}} &\equiv \left\{ \mathbf{y} \in R^{1+n} : y_0 = 1 \text{ and } f_{\mathbf{t}}(\mathbf{y}) \leq 0 \text{ } (\mathbf{t} \in \tilde{T}) \right\}, \\ &= \left\{ \mathbf{y} \in R^{1+n} : \begin{array}{l} y_0 = 1 \text{ and } \mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \mathbf{y} \leq 0 \\ \text{for every } \mathbf{t} \geq \mathbf{0} \text{ such that } \sum_{k=1}^m t_k \mathbf{Q}_k \in \mathcal{S}(n)_+ \end{array} \right\}. \end{aligned} \right\} \quad (4)$$

Although the derivation of the relaxation (3) of the QP (1) is simple and straightforward, it seems difficult to implement the relaxation (3) on computer because the set \tilde{T} over which the index vector \mathbf{t} of the convex-quadratic inequality $f_{\mathbf{t}}(\mathbf{y}) \leq 0$ changes is a continuum, non-polyhedral and convex subset of R^m in general. Under a moderate assumption (Condition 2.2), the main theorem (Theorem 2.3) establishes the equivalence between the SDP relaxation and the relaxation (3) using convex-quadratic valid inequalities. Thus the SDP relaxation may be regarded as an implementable version of the relaxation (3).

We give the main theorem without proof in Section 2. Section 3 states a basic principle which makes it possible for us to extend the Lovász-Schrijver SDP relaxation method for IPs to nonconvex QPs. In Section 4, we present Shor's relaxation method [14], and show some duality relation among the SDP relaxation, Shor's relaxation and the relaxation (3) using convex-quadratic valid inequalities. The discussions in Sections 3 and 4 are not only necessary to prove the main theorem, but also helpful to the readers' deep understanding of the SDP relaxation. Section 5 is devoted to a proof of the main theorem.

2 Main Theorem.

For every $\mathbf{A} \in \mathcal{S}(m)$ and $\mathbf{B} \in \mathcal{S}(m)$, $\mathbf{A} \bullet \mathbf{B}$ denotes their inner product, i.e., $\mathbf{A} \bullet \mathbf{B} \equiv \text{Tr } \mathbf{A}^T \mathbf{B}$ (the trace of $\mathbf{A}^T \mathbf{B}$). It should be noted that any linear function $g : \mathcal{S}(m) \rightarrow R$ can be written as $g(\mathbf{Y}) = \mathbf{A} \bullet \mathbf{Y}$ for some $\mathbf{A} \in \mathcal{S}(m)$. Define

$$\left. \begin{aligned} \mathbf{C} &\equiv \begin{pmatrix} \gamma & \mathbf{d}^T/2 \\ \mathbf{d}/2 & \mathbf{O} \end{pmatrix} \in \mathcal{S}(1+n), \\ \hat{\mathcal{G}} &\equiv \{ \mathbf{Y} \in \mathcal{S}(1+n)_+ : Y_{00} = 1, \mathbf{P}_k \bullet \mathbf{Y} \leq 0 \text{ } (k = 1, 2, \dots, m) \}, \\ \hat{\mathcal{F}} &\equiv \{ \mathbf{Y} \mathbf{e}_0 : \mathbf{Y} \in \hat{\mathcal{G}} \}, \mathbf{e}_0 \equiv \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in R^{1+n}. \end{aligned} \right\} \quad (5)$$

Obviously, $\hat{\mathcal{G}}$ and $\hat{\mathcal{F}}$ are convex subsets of $\mathcal{S}(1+n)$ and R^{1+n} , respectively. We now introduce the SDP which will serve as a relaxation of the QP (1):

$$\text{Minimize } \mathbf{C} \bullet \mathbf{Y} \text{ subject to } \mathbf{Y} \in \hat{\mathcal{G}}; \quad (6)$$

We can rewrite the SDP as a convex minimization problem in the Euclidean space:

$$\text{Minimize } \mathbf{c}^T \mathbf{y} \text{ subject to } \mathbf{y} \in \hat{\mathcal{F}}. \quad (7)$$

The two problems (6) and (7) above are equivalent in the sense that:

Lemma 2.1

1. \mathbf{y} is a feasible solution of the problem (7) if and only if $\mathbf{y} = \mathbf{Y} \mathbf{e}_0$ for some feasible solution \mathbf{Y} of the problem (6).
2. \mathbf{y} is a minimum solution of the problem (7) if and only if $\mathbf{y} = \mathbf{Y} \mathbf{e}_0$ for some minimum solution \mathbf{Y} of the problem (6).
3. $\inf\{\mathbf{C} \bullet \mathbf{Y} : \mathbf{Y} \in \hat{\mathcal{G}}\} = \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\}$. ■

We will be mainly concerned with the convex minimization problem (7) instead of the SDP (6). If we restrict ourselves to QPs derived from 0-1 IPs, our construction of the problem (7) is a special case of the Lovász-Schrijver [11] relaxation method. We impose the following condition on the feasible region $\hat{\mathcal{G}}$ of the SDP (6) in the main theorem below.

Condition 2.2 There is an interior point \mathbf{Y} of the feasible region $\hat{\mathcal{G}}$ of the SDP (6), a $\mathbf{Y} \in \mathcal{S}(1+n)_{++}$ satisfying $Y_{00} = 1$ and $\mathbf{P}_k \bullet \mathbf{Y} < 0$ ($k = 1, 2, \dots, m$). ■

Now we consider the convex minimization problem (3) introduced in the Introduction as a relaxation of the QP (1) using convex-quadratic valid inequalities for \mathcal{F} . If all the extreme points and all the extreme rays of $\tilde{\mathcal{F}}$ are contained in \mathcal{F} then $\tilde{\mathcal{F}}$ coincides with $\text{co } \mathcal{F}$, the convex hull of \mathcal{F} and the problem (3) gives the best convex relaxation of the QP (1). But $\tilde{\mathcal{F}} \neq \text{co } \mathcal{F}$ in general. We focus our attention to a subset of extreme points of $\tilde{\mathcal{F}}$ which are shown to be contained in \mathcal{F} ((iii) of Theorem 2.3). We say that a point $\mathbf{y} \in \tilde{\mathcal{F}}$ is a *strictly convex boundary point* of $\tilde{\mathcal{F}}$ if there exists a $\mathbf{t} = (t_1, t_2, \dots, t_m)^T \geq \mathbf{0}$ such that

$$\mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \mathbf{y} = 0 \text{ and } \sum_{k=1}^m t_k \mathbf{Q}_k \in \mathcal{S}(n)_{++}. \quad (8)$$

It should be noted that the definition of a strictly convex boundary point depends on the algebraic representation of \mathcal{F} . That is, a strictly convex boundary point \mathbf{y} of $\tilde{\mathcal{F}}$ of the representation (4) is not necessarily a strictly convex boundary point of $\tilde{\mathcal{F}}$ of a distinct representation. See section 6.

Now we are ready to state:

Theorem 2.3 (main theorem)

1. $\mathcal{F} \subseteq \hat{\mathcal{F}} \subseteq \tilde{\mathcal{F}}$.
2. Suppose that the feasible region $\hat{\mathcal{G}}$ satisfies Condition 2.2. Then

$$\inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\} = \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \tilde{\mathcal{F}}\} \quad (9)$$
 for every $\mathbf{c} \in R^{1+n}$, and $\tilde{\mathcal{F}} = \text{cl } \hat{\mathcal{F}}$, the closure of $\hat{\mathcal{F}}$.
3. Every strictly convex boundary point \mathbf{y} of $\tilde{\mathcal{F}}$ belongs to \mathcal{F} . ■

Proof of the theorem is given in Section 5.

3 A Single Quadratic Inequality.

The most important principle behind the SDP relaxation is: Any quadratic inequality in the n -dimensional Euclidean space can be relaxed into a linear matrix inequality in $(1+n) \times (1+n)$ symmetric matrices. We will associate each vector $\mathbf{y} = (1, y_1, y_2, \dots, y_n)^T$ in R^{1+n} with a $(1+n) \times (1+n)$ symmetric matrix

$$\mathbf{Y} = \mathbf{y}\mathbf{y}^T = \begin{pmatrix} 1 & y_1 & y_2 & \cdots & y_n \\ y_1 & y_1 y_1 & y_1 y_2 & \cdots & y_1 y_n \\ y_2 & y_2 y_1 & y_2 y_2 & \cdots & y_2 y_n \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ y_n & y_n y_1 & y_n y_2 & \cdots & y_n y_n \end{pmatrix} \in \mathcal{S}(1+n). \quad (10)$$

The matrix $\mathbf{Y} \in \mathcal{S}(1+n)$ contains all the constant, linear and quadratic “atomic” terms, *i.e.*, the nonzero constant term 1, the n linear terms y_1, y_2, \dots, y_n and the n^2 quadratic terms $y_1 y_1, y_1 y_2, \dots, y_n y_n$ in its elements, so that we can represent any function consisting of linear and quadratic forms of y_1, y_2, \dots, y_n in terms of a linear combination of those terms, *i.e.*, a linear function $\mathbf{P} \bullet \mathbf{Y}$ of \mathbf{Y} for some $\mathbf{P} \in \mathcal{S}(1+n)$.

By the construction, for

$$\mathbf{P} \equiv \begin{pmatrix} \pi & \mathbf{q}^T/2 \\ \mathbf{q}/2 & \mathbf{Q} \end{pmatrix} \in \mathcal{S}(1+n), \quad \mathbf{q} \in R^n, \mathbf{Q} \in \mathcal{S}(n), \quad (11)$$

we see that

$$\mathbf{y}^T \mathbf{P} \mathbf{y} = \mathbf{P} \bullet \mathbf{Y} \quad (12)$$

whenever

$$\mathbf{y} = \mathbf{Y} \mathbf{e}_0, \quad y_0 = 1 \quad \text{and} \quad \mathbf{Y} = \mathbf{y}\mathbf{y}^T. \quad (13)$$

On the other hand, we know that an $(1+n) \times (1+n)$ matrix \mathbf{Y} satisfies (13) for some $\mathbf{y} = (y_0, y_1, \dots, y_n)^T \in R^{1+n}$ if and only if

$$\mathbf{y} = \mathbf{Y} \mathbf{e}_0, \quad Y_{00} = 1, \quad \mathbf{Y} \in \mathcal{S}(1+n)_+ \quad \text{and} \quad \text{rank } \mathbf{Y} = 1.$$

Hence

$$\mathbf{y} \in R^{1+n}, \mathbf{y}^T \mathbf{P} \mathbf{y} \leq 0 \text{ and } y_0 = 1$$

if and only if

$$\mathbf{y} = \mathbf{Y} \mathbf{e}_0 \in R^{1+n}, \mathbf{P} \bullet \mathbf{Y} \leq 0, Y_{00} = 1, \mathbf{Y} \in \mathcal{S}(1+n)_+ \text{ and } \text{rank } \mathbf{Y} = 1.$$

Dropping the last rank condition $\text{rank } \mathbf{Y} = 1$, we obtain:

Lemma 3.1 *Let $\mathbf{P} \in \mathcal{S}(1+n)$. If*

$$\mathbf{y}^T \mathbf{P} \mathbf{y} \leq 0 \text{ and } y_0 = 1 \tag{14}$$

then

$$\mathbf{y} = \mathbf{Y} \mathbf{e}_0 \in R^{1+n}, \mathbf{P} \bullet \mathbf{Y} \leq 0, Y_{00} = 1 \text{ and } \mathbf{Y} \in \mathcal{S}(1+n)_+ \tag{15}$$

for some $\mathbf{Y} \in \mathcal{S}(1+n)$. ■

(Relaxation by dropping the rank condition as mentioned above has been utilized in many papers [1, 2, 5, 7, 8, 10, 13, 17], etc.).

Lemma 3.2 *Let \mathbf{P} be a $(1+n) \times (1+n)$ symmetric matrix of the form (11). Suppose that a $(1+n) \times (1+n)$ matrix \mathbf{Y} and $\mathbf{y} \in R^{1+n}$ satisfy the relation (15). Let*

$$\mathbf{y} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \text{ and } \mathbf{Y} = \begin{pmatrix} 1 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{X} \end{pmatrix}.$$

Then

$$\mathbf{y}^T \mathbf{P} \mathbf{y} = \mathbf{P} \bullet \mathbf{Y} - \mathbf{Q} \bullet (\mathbf{X} - \mathbf{x} \mathbf{x}^T) \leq -\mathbf{Q} \bullet (\mathbf{X} - \mathbf{x} \mathbf{x}^T).$$

If in addition the $n \times n$ matrix \mathbf{Q} is positive semi-definite then \mathbf{y} satisfies the relation (14).

Proof: By the definitions of the matrices \mathbf{P} and \mathbf{Y} , we have that

$$\begin{aligned} \mathbf{y}^T \mathbf{P} \mathbf{y} &= \pi + \mathbf{q}^T \mathbf{x} + \mathbf{x}^T \mathbf{Q} \mathbf{x} \\ &= \pi + \mathbf{q}^T \mathbf{x} + \mathbf{Q} \bullet \mathbf{X} - \mathbf{Q} \bullet (\mathbf{X} - \mathbf{x} \mathbf{x}^T) \\ &= \mathbf{P} \bullet \mathbf{Y} - \mathbf{Q} \bullet (\mathbf{X} - \mathbf{x} \mathbf{x}^T) \\ &\leq -\mathbf{Q} \bullet (\mathbf{X} - \mathbf{x} \mathbf{x}^T). \end{aligned}$$

Thus we have shown the first assertion. It follows from $\mathbf{Y} \in \mathcal{S}(1+n)_+$ and $Y_{00} = 1$ that $\mathbf{X} - \mathbf{x} \mathbf{x}^T \in \mathcal{S}(n)_+$. Hence if $\mathbf{Q} \in \mathcal{S}(n)_+$ then $\mathbf{Q} \bullet (\mathbf{X} - \mathbf{x} \mathbf{x}^T) \geq 0$; hence $\mathbf{y}^T \mathbf{P} \mathbf{y} \leq 0$. ■

4 Duality.

Applying Shor's relaxation method [14] to the QP (1), we obtain an SDP

$$\text{Maximize } t_0 \text{ subject to } \mathbf{t} \in T^d, \quad (16)$$

where

$$T^d \equiv \left\{ \mathbf{t} = (t_0, t_1, \dots, t_m)^T : \begin{array}{l} \mathbf{C} - t_0 \mathbf{e}_0 \mathbf{e}_0^T + \sum_{i=1}^m t_i \mathbf{P}_i \in \mathcal{S}(1+n)_+, \\ t_i \geq 0 \ (i = 1, 2, \dots, m) \end{array} \right\}.$$

Between the two problems (16) and (1), the following relation holds.

Lemma 4.1 ([14], see also [17]) *If $\mathbf{t} = (t_0, t_1, \dots, t_m)^T \in R^{1+m}$ is a feasible solution of the SDP (16) and $\mathbf{y} \in R^{1+n}$ a feasible solution of the QP (1), then their objective values t_0 and $\mathbf{c}^T \mathbf{y}$ satisfies the inequality $t_0 \leq \mathbf{c}^T \mathbf{y}$; $\sup\{t_0 : \mathbf{t} \in T^d\} \leq \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \mathcal{F}\}$.*

Proof: Assume that $\mathbf{t} = (t_0, t_1, \dots, t_m)^T \in T^d$ and $\mathbf{y} \in \mathcal{F}$. Then

$$0 \leq \mathbf{y}^T \left(\mathbf{C} - t_0 \mathbf{e}_0 \mathbf{e}_0^T + \sum_{i=1}^m t_i \mathbf{P}_i \right) \mathbf{y} = \mathbf{c}^T \mathbf{y} - t_0 + \sum_{i=1}^m t_i \mathbf{y}^T \mathbf{P}_i \mathbf{y} \leq \mathbf{c}^T \mathbf{y} - t_0.$$

(This proof is essentially due to [17]). ■

The SDP (16) is corresponding to the Lagrangian dual of the QP (1). See the papers [13, 14, 15] for details.

It is easily verified that the SDPs (6) and (16) are dual to each other. Hence, from the duality theorem (see, for example, Theorem 4.2.1 of [12]) and Lemma 2.1, we obtain:

Lemma 4.2 (Duality between (3) and (16))

1. *If $\mathbf{t} = (t_0, t_1, \dots, t_m)^T \in R^{1+m}$ is a feasible solution of the SDP (16) and $\mathbf{y} \in R^{1+n}$ a feasible solution of the problem (3), their objective values t_0 and $\mathbf{c}^T \mathbf{y}$ satisfy $t_0 \leq \mathbf{c}^T \mathbf{y}$; $\sup\{t_0 : \mathbf{t} \in T^d\} \leq \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\}$.*
2. *Suppose that Condition 2.2 holds and that $-\infty < \hat{g} \equiv \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\}$. Then the SDP (16) has a maximum solution $\mathbf{t}^* \in R^{1+m}$ with the maximum objective value $t_0^* = \hat{g}$.* ■

The lemma below establishes a weak duality relation between the convex minimization problem (3) and the SDP (16).

Lemma 4.3 *If $\mathbf{t} = (t_0, t_1, \dots, t_m) \in R^{1+m}$ is a feasible solution of the SDP (16) and $\mathbf{y} \in R^{1+n}$ a feasible solution of the problem (3), their objective values t_0 and $\mathbf{c}^T \mathbf{y}$ satisfy $t_0 \leq \mathbf{c}^T \mathbf{y}$; $\sup\{t_0 : \mathbf{t} \in T^d\} \leq \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \tilde{\mathcal{F}}\}$.*

Proof: Suppose that $\mathbf{t} \in T^d$ and $\mathbf{y} \in R^{1+n} \in \tilde{\mathcal{F}}$. Let

$$\mathbf{Z} \equiv \begin{pmatrix} \zeta & \mathbf{w}^T \\ \mathbf{w} & \mathbf{Q} \end{pmatrix} = \mathbf{C} - t_0 \mathbf{e}_0 \mathbf{e}_0^T + \sum_{k=1}^m t_k \mathbf{P}_k,$$

where $\zeta \in R$, $\mathbf{w} \in R^n$ and $\mathbf{Q} \in \mathcal{S}(n)$. We see by the definitions of the matrices \mathbf{C} , $\mathbf{e}_0 \mathbf{e}_0^T$, $\mathbf{P}_k \in \mathcal{S}(1+n)$ ($k = 1, 2, \dots, m$) that $\mathbf{Q} = \sum_{k=1}^m t_k \mathbf{Q}_k$. On the other hand, it follows from $\mathbf{Z} \in \mathcal{S}(1+n)_+$ that $\mathbf{Q} = \sum_{k=1}^m t_k \mathbf{Q}_k \in \mathcal{S}(n)_+$. Hence we obtain from $\mathbf{y} \in \tilde{\mathcal{F}}$ that $\mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \mathbf{y} \leq 0$. Consequently,

$$0 \leq \mathbf{y}^T \mathbf{Z} \mathbf{y} = \mathbf{y}^T \mathbf{C} \mathbf{y} - t_0 \mathbf{y}^T \mathbf{e}_0 \mathbf{e}_0^T \mathbf{y} + \mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \mathbf{y} \leq \mathbf{c}^T \mathbf{y} - t_0.$$

■

5 Proof of the Main Theorem.

(i) The first inclusion relation $\mathcal{F} \subseteq \hat{\mathcal{F}}$ follows from Lemma 3.1. To prove the second inclusion relation $\hat{\mathcal{F}} \subseteq \tilde{\mathcal{F}}$, assume that $\mathbf{y} \in \hat{\mathcal{F}}$. Then there exists a $\mathbf{Y} \in \hat{\mathcal{G}}$ such that $\mathbf{y} = \mathbf{Y} \mathbf{e}_0$; specifically \mathbf{Y} satisfies $0 \geq \mathbf{P}_k \bullet \mathbf{Y}$ ($k = 1, 2, \dots, m$). Hence

$$\left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \bullet \mathbf{Y} \leq 0 \text{ for every } \mathbf{t} = (t_1, t_2, \dots, t_m)^T \geq \mathbf{0}.$$

By Lemma 3.2, we see that

$$\mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \mathbf{y} \leq 0 \text{ whenever } \sum_{k=1}^m t_k \mathbf{Q}_k \in \mathcal{S}(n)_+.$$

This implies $\mathbf{y} \in \tilde{\mathcal{F}}$. Thus we have shown that $\hat{\mathcal{F}} \subseteq \tilde{\mathcal{F}}$.

(ii) Since $\hat{\mathcal{F}} \subseteq \tilde{\mathcal{F}}$, we know that

$$\inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \tilde{\mathcal{F}}\} \leq \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\} \quad (17)$$

for every $\mathbf{c} \in R^{1+n}$. Let $\mathbf{c} \in R^{1+n}$ be fixed arbitrarily. If $\inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\} = -\infty$ then $\inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \tilde{\mathcal{F}}\} = -\infty$ by (17). Hence we obtain the equality (9). So assume that $\hat{g} \equiv \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\} > -\infty$. By Lemma 4.2, there exists a maximum solution $\mathbf{t}^* = (t_0^*, t_1^*, \dots, t_m^*)^T \in R^{1+m}$ of the SDP (16) with the objective value $t_0^* = \hat{g}$. We also see by Lemma 4.3 that $t_0^* \leq \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \tilde{\mathcal{F}}\}$. Therefore

$$\inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \tilde{\mathcal{F}}\} \leq \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \hat{\mathcal{F}}\} = \hat{g} = t_0^* \leq \inf\{\mathbf{c}^T \mathbf{y} : \mathbf{y} \in \tilde{\mathcal{F}}\}.$$

Thus we have shown the equality (9). By the construction, $\tilde{\mathcal{F}}$ is a closed convex subset of R^{1+n} and $\hat{\mathcal{F}}$ is a convex subset of R^{1+n} . Hence the identity (9) for every $\mathbf{c} \in R^{1+n}$

implies that $\tilde{\mathcal{F}} = \text{cl } \hat{\mathcal{F}}$.

(iii) Assume on the contrary that $\mathbf{y} \notin \mathcal{F}$ for some strictly convex boundary point \mathbf{y} of $\tilde{\mathcal{F}}$. It follows from $\mathbf{y} \notin \mathcal{F}$ that $\mathbf{y}^T \mathbf{P}_j \mathbf{y} > 0$ for some $j \in \{1, 2, \dots, m\}$. Since \mathbf{y} is a strictly convex boundary point of $\tilde{\mathcal{F}}$, there exists some $\mathbf{t} = (t_1, t_2, \dots, t_m)^T \geq 0$ for which

$$y_0 = 1, \mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k \right) \mathbf{y} = 0 \quad \text{and} \quad \sum_{k=1}^m t_k \mathbf{Q}_k \in \mathcal{S}(n)_{++}$$

holds. Hence if $\epsilon > 0$ is sufficiently small, we obtain

$$y_0 = 1, \mathbf{y}^T \left(\sum_{k=1}^m t_k \mathbf{P}_k + \epsilon \mathbf{P}_j \right) \mathbf{y} > 0 \quad \text{and} \quad \sum_{k=1}^m t_k \mathbf{Q}_k + \epsilon \mathbf{Q}_j \in \mathcal{S}(n)_{++},$$

which is a contradiction to the assumption that $\mathbf{y} \in \tilde{\mathcal{F}}$. This completes the proof of the main theorem.

6 Concluding Discussion.

The effectiveness of the SDP relaxation for a nonconvex QP (or a 0-1 IP) depends on the representation of its feasible region using linear and/or quadratic inequalities. Suppose that the feasible region \mathcal{F} of the canonical form QP (1) is bounded and involves some linear inequality constraints

$$\mathbf{y}^T \mathbf{P}_k \mathbf{y} \equiv \pi_k + \mathbf{q}_k^T \mathbf{x} \leq 0 \quad (k \in K),$$

where $\mathbf{y} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix}$, and $K \subseteq \{1, 2, \dots, m\}$. Let S denote the polyhedral region determined by these linear inequalities;

$$S = \left\{ \mathbf{y} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} : \pi_k + \mathbf{q}_k^T \mathbf{x} \leq 0 \quad (k \in K) \right\}.$$

We want to cut off all the vertices of S that do not lie in \mathcal{F} when we apply the SDP relaxation. As we will see below, this is always possible if we replace those linear inequality constraints by convex-quadratic inequality constraints

$$\mathbf{y}^T \mathbf{P}'_k \mathbf{y} \equiv (\pi_k + \mathbf{q}_k^T \mathbf{x})(\pi'_k + \mathbf{q}_k^T \mathbf{x}) \leq 0 \quad (k \in K),$$

where

$$\mathbf{P}'_k \equiv \begin{pmatrix} \pi'_k \pi_k & (\pi'_k + \pi_k) \mathbf{q}_k^T / 2 \\ (\pi'_k + \pi_k) \mathbf{q}_k / 2 & \mathbf{Q}'_k \end{pmatrix} \in \mathcal{S}(1+n), \quad \mathbf{Q}'_k \equiv \mathbf{q}_k \mathbf{q}_k^T \in \mathcal{S}(n)_+ \quad (k \in K),$$

and π'_k ($k \in K$) are sufficiently large numbers such that

$$\pi_k + \mathbf{q}_k^T \mathbf{x} \leq 0 \leq \pi'_k + \mathbf{q}_k^T \mathbf{x} \quad (k \in K) \quad \text{for every } \mathbf{y} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix} \in \mathcal{F}. \quad (18)$$

Let

$$\mathcal{F}' = \left\{ \mathbf{y} \in R^{1+n} : \begin{array}{l} y_0 = 1, \\ \mathbf{y}^T \mathbf{P}_i \mathbf{y} \leq 0 \ (i \in \{1, 2, \dots, m\} \setminus K), \\ \mathbf{y}^T \mathbf{P}'_k \mathbf{y} \leq 0 \ (k \in K) \end{array} \right\}$$

The condition (18) above on π'_k ($k \in K$) ensures that $\mathcal{F}' = \mathcal{F}$. Suppose that $\mathbf{y} = \begin{pmatrix} 1 \\ \mathbf{x} \end{pmatrix}$ is a vertex of S . Then there exists a subset K' of K such that

$$\mathbf{y}^T \mathbf{P}_k \mathbf{y} = \pi_k + \mathbf{q}_k^T \mathbf{x} = 0 \ (k \in K') \text{ and } \{\mathbf{q}_k \in R^n : k \in K'\} \text{ forms a basis of } R^n.$$

Hence

$$\sum_{k \in K'} \mathbf{y}^T \mathbf{P}'_k \mathbf{y} = 0 \text{ and } \sum_{k \in K'} \mathbf{Q}'_k \in \mathcal{S}(n)_{++}.$$

If in addition $\mathbf{y} \in \tilde{\mathcal{F}}'$ then \mathbf{y} is a strictly convex boundary point of $\tilde{\mathcal{F}}'$; hence $\mathbf{y} \in \mathcal{F}'$ by (iii) of Theorem 2.3. Therefore we can conclude that every vertex \mathbf{y} of S belongs to $\tilde{\mathcal{F}}'$ if and only if $\mathbf{y} \in \mathcal{F}'$.

Acknowledgment. The authors are grateful to Professors Stephen Boyd, Farid Alizadeh, Henry Wolkowicz for their helpful comments and references. In particular, the paper [14] was brought to the authors by Stephen Boyd.

References

- [1] W. F. Alizadeh, "Interior point methods in semidefinite programming with application to combinatorial optimization," *SIAM Journal on Optimization* **5** (1995) 13–51.
- [2] W. F. Alizadeh, J. -P. A. Haeberly and M. L. Overton, "Primal-dual interior-point methods for semidefinite programming," 1994.
- [3] S. Boyd, L. E. Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, (SIAM, Philadelphia, 1994).
- [4] R. M. Freund, "Complexity of an algorithm for finding an approximate solution of a semidefinite program with no regularity assumption," Technical report OR 302-94, Operations Research Center, MIT, 1994.
- [5] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of Assoc. Comput. Mach.* to appear. . A preliminary version appeared in *Proceedings of the 26th Annual ACM Symposium on Theory of Computing* (1994) 422–431.

- [6] C. Helmberg, S. Poljak, F. Rendl and H. Wolkowicz, "Combining semidefinite and polyhedral relaxation for integer programs," Lecture note in Computer Science 538 (1995) 124–134.
- [7] C. Helmberg, F. Rendl, R. J. Vanderbei and H. Wolkowicz, "An interior-point method for semidefinite programming," SIAM Journal on Optimization, to appear.
- [8] D. E. Knuth, "The sandwich theorem," *Electronic Journal of Combinatorics* 1 (1995) 1–48.
- [9] M. Kojima, S. Shindoh and S. Hara, "Interior-point methods for the monotone semidefinite linear complementarity problems," Research Report #282, Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152, Japan, April 1994, Revised April 1995.
- [10] L. Lovász, "On the Shannon capacity of a graph," *IEEE Transactions on Information Theory* 25 (1979) 1–7.
- [11] L. Lovász and A. Schrijver, "Cones of matrices and set functions and 0-1 optimization," *SIAM Journal on Optimization* 1 (1991) 166–190.
- [12] Ju. E. Nesterov and A. S. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications* (SIAM, Philadelphia, 1993).
- [13] S. Poljak, F. Rendl and H. Wolkowicz, "A recipe for semidefinite relaxation for (0,1)-quadratic programming," *Journal of Global Optimization* 7 (1995) 51–73.
- [14] N. Z. Shor, "Quadratic optimization problems," *Soviet Journal of Computer and Systems Sciences* 25 (1987) 1–11.
- [15] N. Z. Shor, "Dual quadratic estimates in polynomial and boolean programming," *Annals of Operations Research* 25 (1990) 163–168.
- [16] L. Vandenberghe and S. Boyd, "A primal-dual potential reduction method for problems involving matrix inequalities," *Mathematical Programming* 69 (1995) 205–236.
- [17] L. Vandenberghe and S. Boyd, "Semidefinite Programming," Informations Systems Laboratory, Stanford University, 1994.